

Extensiones del modelo

Emilio Congregado
Universidad de Huelva

3

Sesión 3

- 1.- Multicolinealidad
- 2.- Errores no esféricos: heteroscedasticidad y autocorrelación
- 3.- Heteroscedasticidad
- 4.- Autocorrelación
- 5.- Normalidad

1.- Multicolinealidad

Si las variables explicativas están correlacionadas entre sí, la matriz $(X'X)$ no es invertible, por lo que no podremos calcular el estimador OLS.

El caso más habitual se suele dar cuando construimos una dummy de país e incluimos una constante. En este caso, la constante dependerá linealmente de estas variables ficticias. Pero este problema es fácil de resolver eliminando la constante. El problema al que nos referimos exactamente no es a éste sino a aquéllos casos en los que existe una cierta dependencia lineal, medida ésta a través de la existencia de una cierta correlación lineal entre las variables. En estos casos, la matriz $(X'X)$ se podrá invertir pero los estimadores OLS serán bastante imprecisos.

Comencemos por el caso de multicolinealidad perfecta.

Sea el modelo:

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

Supongamos que la variable x_1 se puede escribir como una combinación lineal de las variables x_2 y x_3 . En este caso podríamos reescribir el modelo inicial como:

$$y = \beta_1(x_2 + x_3) + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

Por lo que el modelo quedaría como:

$$y = (\beta_1 + \beta_2)x_2 + (\beta_1 + \beta_3)x_3 + \varepsilon$$

Por lo que haciendo un cambio de variables:

$$y = \underbrace{(\beta_1 + \beta_2)}_{\gamma} x_2 + \underbrace{(\beta_1 + \beta_3)}_{\psi} x_3 + \varepsilon, \text{ el modelo se limita a estimar } \gamma \text{ y } \psi.$$

Pero si observamos las ecuaciones tenemos tres incógnitas y dos ecuaciones. El modelo está subidentificado, es decir, el problema al que nos enfrentamos es un problema de identificación.

Este no es el problema habitual. El problema que se nos suele plantear es que la correlación entre dos variables explicativas es alta, pero no perfecta. ¿Cuáles son los síntomas?

- 1.- Pequeños cambios en los datos, o la agregación/eliminación de alguna variable causa grandes cambios en los parámetros estimados. La causa está en que $(X^t X)$ es una matriz casi singular, es decir con determinante cercano a 0.
- 2.- Los coeficientes estimados son muy poco significativos, con t's muy bajas, pero la significatividad conjunta es muy alta.
- 3.- La magnitud y el signo de los coeficientes estimados suelen ser raros.

¿Cómo detectar la multicolinealidad?

Si sospechamos que una de las variables explicativas está correlacionada con algunas de las demás, podemos regresarlas y ver si su coeficiente de determinación está cercano a la unidad.

Soluciones a la multicolinealidad

- 1.- Omitir la variable o variables que causan problemas
- 2.- Componentes principales. Este es el método más habitual. Consiste en encontrar una serie de proyecciones ortogonales, un cambio de base que garantice la ortogonalidad de las nuevas variables. El problema asociado al uso de este método (abstracción hecha de la falta de insesgadez del estimador) viene dado por la dificultad a la hora de interpretar los coeficientes.

2.- Errores no esféricos

Recordemos que uno de los supuestos del modelo es que la matriz de varianzas y covarianzas del error es la matriz identidad multiplicada por una constante:

$$\text{var}(\varepsilon) = \begin{pmatrix} \sigma^2 & 0 & \dots & \dots & 0 \\ 0 & \sigma^2 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \sigma^2 \end{pmatrix} = \sigma^2 I$$

Es decir, que todas las covarianzas son nulas y que las varianzas son constantes.

Supongamos por el contrario que la matriz de varianzas y covarianzas fuera de la forma:

$$\text{var}(\varepsilon) = \sigma^2 \Omega$$

Cuando se incumple el hecho de que las varianzas sean constantes pero las covarianzas son nulas, decimos que el modelo presenta **heteroscedasticidad**.

Cuando las varianzas son constantes pero las covarianzas no son nulas, decimos que el modelo presenta **autocorrelación**.

3.- Heteroscedasticidad

En general este problema suele aparecer cuando la variable dependiente varía con la escala del problema. En este caso la matriz de varianzas y covarianzas del error es:

$$\text{var}(\varepsilon) = \sigma^2 \Omega = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}$$

Donde, puede apreciar que no existe correlación entre los residuos.

Consecuencias de la heteroscedasticidad:

- 1.- El estimador sigue siendo insesgado
- 2.- La matriz de varianzas y covarianzas del estimador ya no será $\text{var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$, sino que ahora tendremos:

$$\text{var}(\hat{\beta}) = \sigma^2 (X'X)^{-1} (X'\Omega X) (X'X)$$

Por tanto, $\hat{\sigma}^2 = \frac{e'e}{n-k}$ ya no es un estimador insesgado de σ^2 , por lo que los contrastes utilizados ya no son válidos.

Estimación: Mínimos Cuadrados generalizados

Para tener un modelo que vuelva a cumplir las propiedades del modelo clásico de regresión, se transforma el modelo original. La solución es lo que se llama mínimos cuadrados generalizados.

4.- Autocorrelación

En esta sección vamos a repasar los contrastes de autocorrelación más frecuentes que aparecen en los paquetes econométricos: el contraste Durbin-Watson y el test de los multiplicadores de Lagrange (test de Breusch-Godfrey).

4.1. Test de Durbin Watson

El contraste que se usa tradicionalmente para testar la presencia de autocorrelación, es el propuesto por Durbin y Watson (DW) que viene dado por,

$$DW = d = \frac{\sum_{t=2}^N (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=2}^N \hat{e}_t^2}$$

donde $\hat{e}_t, t = 1, \dots, n$, son los residuos obtenidos por OLS¹.

Consideremos el modelo:

¹ Este test también se utiliza como indicador de una mala especificación de los modelos.

$$y_t = \beta_1 + \beta_2 x_{1t} + \dots + \beta_k x_{kt} + u_t$$

Con

$$u_t \approx N(0, \sigma^2)$$

El contraste de Durbin Watson se utiliza para contrastar si existe correlación de primer orden, es decir, si los errores están generados por un AR(1).

$$u_t = \rho u_{t-1} + \varepsilon_t, t = 1, \dots, N$$

En este contexto, la hipótesis nula de no presencia de autocorrelación $\rho = 0$ tiene como hipótesis alternativas la existencia de autocorrelación positiva $\rho > 0$ - o negativa $\rho < 0$ -.

La tabla de valores críticos del test de Durbin Watson nos da los límites inferiores y superiores para los diferentes valores de n y k .

Entre estos límites existe un valor c , tal que:

$P(d < c) = 0,05$ sobre la hipótesis nula de inexistencia de autocorrelación. ¿Cómo se aplica el test:

1. Se estima el modelo por OLS y se obtienen los residuos:

$$\hat{e}_t = y_t - \hat{\beta}_1 - \hat{\beta}_2 x_{1t} - \dots - \hat{\beta}_k x_{kt}$$

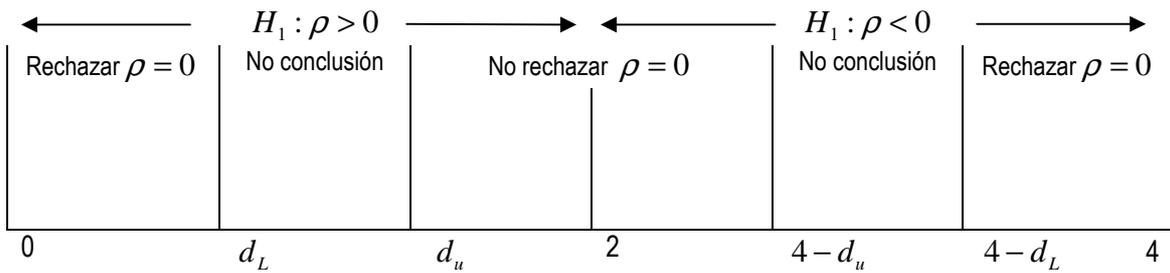
2. Se calcula el valor del test:

$$DW = d = \frac{\sum_{t=2}^N (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=2}^N \hat{e}_t^2} \quad (\text{recuerde que } d, \text{ ha de estar entre } 0 \text{ y } 4 \text{ y que } k, \text{ es el}$$

número de coeficientes excluyendo la constante).

3.- Para contrastar $H_0 : \rho = 0$ frente a $H_1 : \rho > 0$ (autocorrelación positiva) utilizamos un test unitalateral, es decir, vemos en la tabla los valores críticos de d_L y d_u . Si $d \leq d_L$ se rechaza la hipótesis nula. Si $d \geq d_u$ no rechazamos la hipótesis nula, mientras que si $d_L < d < d_u$ el contraste no es concluyente.

4.- Para contrastar $H_0 : \rho = 0$ frente a $H_1 : \rho < 0$ (autocorrelación negativa) utilizamos un test unitalateral, es decir, vemos en la tabla los valores críticos de d_L y d_u . Si $4 - d \leq d_L$ se rechaza la hipótesis nula, es decir, existe autocorrelación negativa. Si $4 - d \geq d_u$ no rechazamos la hipótesis nula (es decir, no existe autocorrelación negativa), mientras que si $d_L < 4 - d < d_u$ el contraste no es concluyente.



4.2. Test LM (test de Breusch-Godfrey)

Consideremos el modelo:

$$y_t = \alpha_1 y_{t-1} + \beta_0 x_{0t} + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t$$

Este modelo se puede representar como:

$$\left. \begin{aligned} y_t &= \alpha_1 y_{t-1} + \beta_0 x_t + u_t \\ u_t &= \rho u_{t-1} + \varepsilon_t \end{aligned} \right\} y_t = \alpha_1 y_{t-1} + \beta_0 x_t + u_t = \alpha_1 y_{t-1} + \beta_0 x_t + \rho u_{t-1} + \varepsilon_t =$$

$$= \alpha_1 y_{t-1} + \beta_0 x_t + \rho(y_{t-1} - \alpha_1 y_{t-2} - \beta_0 x_{t-1}) + \varepsilon_t$$

El test LM puede considerarse como un test de omisión de variables, es decir de exclusión de u_{t-1} de la regresión:

$$\hat{u}_t = \alpha_1 y_{t-1} + \beta_0 x_t + \gamma \hat{u}_t + v_t$$

El test LM puede generalizarse para contrastar la autocorrelación de orden superior, incluyendo en la ecuación anterior un número mayor de retardos de la variable u, de acuerdo con el tipo de autocorrelación que queramos contrastar.

Así, para contrastar que u, es generado por un AR(p), y considerando la hipótesis nula como:

$H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0$, este test sigue una distribución $LM = nR^2 \approx \chi_p^2$, test conocido como test de Breusch-Godfrey.

5.- Normalidad

Algunas de las propiedades del estimador MCO dependen del supuesto de normalidad del término del error. Jarque y Bera proponen el siguiente estadístico para testar la normalidad, para los errores:

$$BJ = T \left(\frac{(asimetría)^2}{6} + \frac{(curtosis - 3)^2}{24} \right) \approx \chi^2(2), \text{ donde la hipótesis nula es la de normalidad.}$$

La hipótesis nula se acepta si: $BJ < \chi^2(2)$